UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/607,370	06/30/2000	Reiner Kraft	ARC9-2000-0046-US1	8705

23334 7590 11/21/2006

FLEIT, KAIN, GIBBONS, GUTMAN, BONGINI
& BIANCO P.L.
ONE BOCA COMMERCE CENTER
551 NORTHWEST 77TH STREET, SUITE 111
BOCA RATON, FL 33487

EXAMINER	
STORK, KYLE R	

ART UNIT	PAPER NUMBER
2178	

DATE MAILED: 11/21/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

UNITED STATES PATENT AND TRADEMARK OFFICE

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

BEFORE THE BOARD OF PATENT APPEALS AND INTERFERENCES

Application Number: 09/607,370
Filing Date: June 30, 2000
Appellant(s): KRAFT ET AL.

Jon Gibbons, Reg. No. 37,333
For Appellant

EXAMINER'S ANSWER

This is in response to the appeal brief filed 7 September 2006 appealing from the Office action mailed 15 July 2005.

(1) Real Party in Interest

A statement identifying by name the real party in interest is contained in the brief.

(2) Related Appeals and Interferences

The examiner is not aware of any related appeals, interferences, or judicial proceedings which will directly affect or be directly affected by or have a bearing on the Board's decision in the pending appeal.

(3) Status of Claims

The statement of the status of claims contained in the brief is correct.

(4) Status of Amendments After Final

The appellant's statement of the status of amendments after final rejection contained in the brief is correct.

(5) Summary of Claimed Subject Matter

The summary of claimed subject matter contained in the brief is correct.

(6) Grounds of Rejection to be Reviewed on Appeal

The appellant's statement of the grounds of rejection to be reviewed on appeal is correct.

(7) Claims Appendix

The copy of the appealed claims contained in the Appendix to the brief is correct.

(8) Evidence Relied Upon

6638314	Mererzon et al.	10-2003
6289342	Lawrence et al.	9-2001

6026409	Blumenthal	2-2000
6523022	Hobbs	2-2003

(9) Grounds of Rejection

The following ground(s) of rejection are applicable to the appealed claims:

Claim Rejections - 35 USC § 103

4. **The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:**

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

5. **Claims 1-3, 14-16, 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Meyerzon et al. (herein after Meyerzon) U.S. Patent No. 6,638,314 B1 filed 6/26/1998 in view of Lawrence et al. (herein after Lawrence) U.S. Patent No. 6,289,342 B1 filed 5/20/1998 and in further view of Blumenthal U.S. Patent No. 6,026,409 filed 9/26/1996.**

In regard to independent claim 1, Meyerzon discloses *retrieving a web document at an address and extracting contents of the web document for rendering an intermediate dynamically constructed in-memory web page representation of the web document at a hub processing unit which is formatted as if displayed for viewing on an end-user's web browser* (Meyerzon Col 7 Lines 60-65 and Col 8 Lines 15-20 i.e. web crawler program searches remote server computers connected to the network for electronic documents and retrieves electronic documents and associated data and a browser displays documents to a user)*; loading secondary documents associated with*

the web document in order to render the secondary documents as part of the in-memory web page representation (Meyerzon Col 8 Lines 26-35 i.e. the client computer transmits data to a search engine, the search engine examines its associated index to find documents and returns the documents which are secondary documents and lists the documents for the user to view), *wherein the secondary documents include one or more images with textual content embedded therein* (Meyerzon Col 9 Lines 44-50 i.e. visual element m include text and hyperlink to an image)*; analyzing and summarizing the in-memory web page representation to produce a text map for the web page document of the textual contents* (Meyerzon Col 10 Lines 13-16 i.e. passes the lists of properties and text to the indexing engine and the indexing engine creates an index, which is used by the search engine in subsequent searches).

Meyerzon does not specifically mention *using optical character recognition* on the images to extract textual content for adding to the textual map for the web page document. However, Lawrence mentions extracting data using optical character recognition (Lawrence Col 7 Lines 51-56 i.e. conversion to electronic form by use of OCR). It would have been obvious to one of ordinary skill in the art at the time of the invention, to apply Lawrence to Meyerzon, providing Meyerzon the benefit of extracting content from a document using OCR, which is quicker the typing out an entire document manually by hand.

Meyerzon does not specifically mention *wherein the hub processing unit renders the in-memory webpage prior to analyzing and summarizing the in-memory webpage.* However, Blumenthal mentions a document that can be rendered prior to user actions

(Blumenthal Col 17 Lines 45-53). It would have been obvious to one of ordinary skill in the art at the time of the invention, to apply Blumenthal to Meyerzon, providing Meyerzon the benefit of rendering the document prior to the user action to ensure the correct page is being analyzed and summarized.

In regard to dependent claim 2, which depends on claim 1, Meyerzon discloses *wherein the retrieving the web document at an address further comprises retrieving a document at an address selected from the group of addresses consisting of a nodal address, a network address, a URL and equivalents* (Meyerzon Col 21 Lines 1-11 i.e. a request to retrieve a list of electronic documents and retrieving a set of document address specifications corresponding the electronic documents).

In regard to dependent claim 3, which depends on claim 1, Meyerzon discloses *wherein the one or more images with textual content embedded therein include at least one of an in-line GIF image and an in-line JPEG image.* (Meyerzon Col 9 Lines 37-46 i.e. an image is retrieved to display on a web page and it is well known in the art the images displayed on web pages can be a gif and jpeg image).

In regard to dependent claim 7, which depends on claim 1, Meyerzon discloses *initializing a first list with seed values* (Meyerzon Col 17 Lines 25-26 i.e. assigning a current crawl number to the current web crawl)*; checking if there are any URLs to be processed and in response that any URL exists to be processed then performing the follow sub-steps of* (Meyerzon Col 17 Lines 28-29 i.e. determine whether an electronic document has been retrieved)*: determining if a URL is in a second list; and in response that a URL is not in the second list then performing the following sub-*

steps of: inserting the URL into the first list; scheduling the URL for crawling; crawling the URL when scheduled to do so; removing the URL from the first list after the scheduled crawling; entering the URL into the second list (Meyerzon Col 9 Lines 64 and Col 10 Lines 1-11 i.e. history map checks each hyperlink URL to determine if it is already listed in the history map, if not the URLs are added and are marked as not being crawled and added to the transaction log. The history map includes a number crawled and number modified data)*; and repeating the checking step until there are no more URLs to be processed; where if the determining step determines that the URL is in the second list then repeating the checking step until there are no more URLs to be processed.* (Meyerzon Col 12 Lines 1-17 i.e. retrieves and processed a URL until there are none left in the transaction log)

In regard to dependent claim 8, which depends on claim 7, Meyerzon discloses *wherein the sub-step of initializing a first list with seed values further includes the list being a URL pool.* (Meyerzon Col 7 Lines 65-67 i.e. retrieving a processing URLs from the transaction log)

In regard to dependent claim 9, which depends on claim 7, Meyerzon discloses *wherein the sub-step of determining if a URL is in a second list further includes the second list being a visited pool.* (Meyerzon Figure 4 shows a column indicating the number crawled and modified)

In regard to dependent claim 10, which depends on claim 7, Meyerzon discloses *wherein the sub-step of crawling further comprises the sub-steps of: issuing an HTTP command to a web server named in the URL; receiving contents of an HTML*

page as a result of the issued HTTP command; and passing on the contents of the HTML page to a Page Rendering subroutine. (Meyerzon Col 8 Lines 26-35 i.e. the client computer transmits data to a search engine, the search engine examines its associated index to find documents and returns the documents which are secondary documents and lists the documents for the user to view)

In regard to dependent claim 11, which depends on claim 10, Meyerzon discloses *receiving the contents of the HTML page in the Page Rendering subroutine; building an in-memory representation of a layout for the HTML page and if more data is needed to properly form the representation, then performing the sub-steps of* (Meyerzon Col 7 Lines 60-65 and Col 8 Lines 15-20 i.e. web crawler program searches remote server computers connected to the network for electronic documents and retrieves electronic documents and associated data and a browser displays documents to a user)*: requesting additional web-based information; gathering this additional web-based information; inserting any URLs associated with this additional web-based information into the second list and a URL cache* (Meyerzon Col 9 Lines 37-46 i.e. an image is retrieved to display on a web page)*; building a final amended representation; and forwarding the final amended representation to an Extraction subroutine; wherein, if no more data is needed to properly form the in-memory representation, then forwarding the in-memory representation to the Extraction subroutine.* (Meyerzon Col 16 Lines 32-44)

In regard to dependent claim 12, which depends on claim 11, Meyerzon discloses *accessing a set of memory structures of the Page Renderer* (Meyerzon Col 6 Lines 23-60 i.e. accessing local and remote memory devices)*; copying a text portion of*

the structures into a text map (Meyerzon Col 15 Lines 15-16 i.e. copying all of the history map entries into the transaction log as entries); *inspecting any in-line GIF and JPEG image references in the memory structures* (Meyerzon Col 9 Lines 37-46 i.e. an image is retrieved to display on a web page and it is well known in the art the images displayed on web pages can be a gif and jpeg image); *extracting alternate text attributes* (Meyerzon Col 5 Lines 7-8 i.e. extracting data from each of the retrieved documents); *adding the alternate text attributes to a text map* (Meyerzon Col 2 Lines 48-51 i.e. information from the electronic document retrieved from the web crawl is stored in an index); *extracting text content from the GIF and JPEG images; adding text content from the images to the text map* (Meyerzon Col 9 Lines 37-46 i.e. an image is retrieved to display on a web page and it is well known in the art the images displayed on web pages can be a gif and jpeg image Col 5 Lines 7-8 i.e. extracting data from each of the retrieved documents Col 2 Lines 48-51 i.e. information from the electronic document retrieved from the web crawl is stored in an index); *and forwarding the text map to a Page Summarizer subroutine.* (Meyerzon Col 9 Lines 64 and Col 10 Lines 1-11 i.e. history map checks each hyperlink URL to determine if it is already listed in the history map, if not the URLs are added and are marked as not being crawled and added to the transaction log. The history map includes a number crawled and number modified data)

Meyerzon does not specifically mention *invoking an optical character recognition engine;* analyzing any in-line GIF and JPEG images *using the optical character recognition* engine for text content. However, Lawrence mentions extracting data using optical character recognition (Lawrence Col 7 Lines 51-56 i.e. conversion to electronic

form by use of OCR). It would have been obvious to one of ordinary skill in the art at the time of the invention, to apply Lawrence to Meyerzon, providing Meyerzon the benefit of extracting content from a document using OCR, which is quicker the typing out an entire document manually by hand.

In regard to dependent claim 13, which depends on claim 12, Meyerzon discloses *receiving a text map from the Page Extractor subroutine; processing the text map in an application-specific manner* (Meyerzon Col 2 Lines 48-51 i.e. information from the electronic document retrieved from the web crawl is stored in an index to begin the routine)*; applying data extraction patterns to the text map* (Meyerzon Col 5 Lines 7-8 i.e. extracting data from each of the retrieved documents)*; translating resultant data from the applying step; forwarding any URLs present in the text map to a manager subroutine; and forwarding any extracted data and metadata to application logic.* (Meyerzon Col 9 Lines 64 and Col 10 Lines 1-11 i.e. history map checks each hyperlink URL to determine if it is already listed in the history map, if not the URLs are added and are marked as not being crawled and added to the transaction log. The history map includes a number crawled and number modified data)

In regard to independent claims 14 and 20, claims 14 and 20 in addition to the following reflect similar subject matter claimed in claim 1 and are rejected along the same rationale. (Meyerzon Col 20 Lines 13-14 i.e. computer readable medium having computer executable instruction and Col 20 Lines 23-24 i.e. a system for retrieving stored information)

In regard to dependent claim 15, which depends on claim 14, claim 15 in addition to the following reflect similar subject matter claimed in claim 2 and are rejected along the same rationale. (Meyerzon Col 20 Lines 13-14 i.e. computer readable medium having computer executable instruction)

In regard to dependent claim 16, which depends on claim 14, claim 16 in addition to the following reflect similar subject matter claimed in claim 3 and are rejected along the same rationale. (Meyerzon Col 20 Lines 13-14 i.e. computer readable medium having computer executable instruction)

6. **Claims 4-6 and 17-19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Meyerzon et al. (herein after Meyerzon) in view of Lawrence et al. (herein after Lawrence) as applied to claim 1 and in further view of Hobbs U.S. Patent No. 6,523,022 B1 filed 7/7/1999.**

In regard to dependent claim 4, which depends on claim 1, Meyerzon does not specifically mention *wherein the loading secondary documents further comprises the loading of secondary documents including one or more Java applets with textual content embedded therein.* However, Hobbs mentions that Java applets are used (Hobbs Col 28 Line 35). It would have been obvious to one of ordinary skill in the art at the time of the invention, to apply Hobbs to Meyerzon, providing Meyerzon the benefit of using Java Applets for web pages in the process of searching the web documents because Java Applets are compatible with many web pages and browsers.

In regard to dependent claim 5, which depends on claim 1, Meyerzon does not specifically mention *wherein the loading secondary documents further comprises the loading of secondary documents including web documents selected from the group of documents consisting of in-line frames, frames, and equivalents.* However, Hobbs mentions that frames and in-line frames are used (Hobbs Col 7 Lines 63 through Col 8 Lines 1-34). It would have been obvious to one of ordinary skill in the art at the time of the invention, to apply Hobbs to Meyerzon, providing Meyerzon the benefit of using frames and in-line frames for easy viewing for the user.

In regard to dependent claim 6, which depends on claim 4, Meyerzon does not specifically mention *wherein the loading secondary documents further comprises the loading of secondary documents including one or more Java Script components with textual content embedded therein.* However, Hobbs mentions that Java applets are used (Hobbs Col 28 Line 35). It would have been obvious to one of ordinary skill in the art at the time of the invention, to apply Hobbs to Meyerzon, providing Meyerzon the benefit of using Java Scripts for web pages in the process of searching the web documents because Java Scripts are compatible with many web pages and browsers.

In regard to dependent claim 17, which depends on claim 14, claim 17 in addition to the following reflect similar subject matter claimed in claim 4 and are rejected along the same rationale. (Meyerzon Col 20 Lines 13-14 i.e. computer readable medium having computer executable instruction)

In regard to dependent claim 18, which depends on claim 14, claim 18 in addition to the following reflect similar subject matter claimed in claim 5 and are rejected

along the same rationale. (Meyerzon Col 20 Lines 13-14 i.e. computer readable medium having computer executable instruction)

In regard to dependent claim 19, which depends on claim 17, claim 19 in addition to the following reflect similar subject matter claimed in claim 6 and are rejected along the same rationale. (Meyerzon Col 20 Lines 13-14 i.e. computer readable medium having computer executable instruction)

(10) Response to Argument

The appellant begins by arguing that Meyerzon, Lawrence, and Blumenthal fail to teach or suggest, "rendering an intermediate dynamically constructed in-memory web page representation of the web document at a hub processing unit which is formatted as if displayed for viewing on an end user's web browser (pages 7-8)." The examiner respectfully disagrees. As the appellant points out, rendering "is not implying a visual display of a document, but rather the construction of a data structure of the webpage in memory, which is subsequently analyzed and summarized (Appeal Brief: page 17, first full paragraph)." Although it is unclear precisely how the document is rendered as if displayed for viewing on an end user's web browser merely by construction a data structure of the webpage in memory, it is irrelevant. As the appellant admits, Meyerzon discloses a gatherer, extractor, and summarizer at a hub processing unit (Appeal Brief: page 12, diagram). The diagram fails to show rendering at the hub processing unit.

However, Meyerzon discloses **retrieving a document** and a filter daemon **parsing the electronic document** (column 9, lines 29-31). First, retrieving a document

to the hub, would construct a data structure of the webpage in memory, as simply storing a document in memory generates the data structure holding the document. However, upon parsing the retrieved electronic document, a list of text and properties is generated (column 9, lines 29-31). The text is the text that is to be displayed in a web browser program and metadata describing the formatting of the text for display in the browser (column 9, lines 33-35).

The appellant further relies upon this limitation as the reason that the second and third limitations of claim 1 cannot be valid (page 14). However, the examiner respectfully disagrees. Since the limitation is taught by Meyerzon, this argument is non-persuasive.

Further, the appellant argues that Lawrence fails to teach loading secondary documents associated with the web document (page 14). However, Lawrence is not relied upon to teach this limitation. Meyerzon discloses loading secondary document associated with the web document (column 8, lines 26-35).

With respect to Blumenthal, the appellant argues that Blumenthal fails to teach the web processing unit rendering the in-memory webpage prior to analyzing and summarizing the in-memory webpage (page 15). This is based upon the applicant's statement that rendering "is not implying a visual display of a document, but rather the construction of a data structure of the webpage in memory, which is subsequently analyzed and summarized (Appeal Brief: page 17, first full paragraph)." First, in order to render a document for display, Blumenthal must first construct a data structure of the

webpage in memory. Further, Meyerzon discloses the construction of a webpage in memory (column 9, lines 29-31).

Finally, the appellant appears to argue that the examiner's conclusion of obviousness is based upon improper hindsight reasoning (page 17). However, it must be recognized that any judgment on obviousness is in a sense necessarily a reconstruction based upon hindsight reasoning. But so long as it takes into account only knowledge which was within the level of ordinary skill at the time the claimed invention was made, and does not include knowledge gleaned only from the applicant's disclosure, such a reconstruction is proper. See *In re McLaughlin*, 443 F.2d 1392, 170 USPQ 209 (CCPA 1971). In addition, the appellant argues that the incorporation of the references would destroy the intent, purpose, and/or function of the references (page 18). However, the test for obviousness is not whether the features of a secondary reference may be bodily incorporated into the structure of the primary reference; nor is it that the claimed invention must be expressly suggested in any one or all of the references. Rather, the test is what the combined teachings of the references would have suggested to those of ordinary skill in the art. See *In re Keller*, 642 F.2d 413, 208 USPQ 871 (CCPA 1981).

(11) Related Proceeding(s) Appendix

No decision rendered by a court or the Board is identified by the examiner in the Related Appeals and Interferences section of this examiner's answer.

For the above reasons, it is believed that the rejections should be sustained.

Respectfully submitted,

krs

STEPHEN HONG
SUPERVISORY PATENT EXAMINER

Conferees:

Stephen Hong, SPE 2178

Heather Herndon, SPE 2174